

Figure 1A

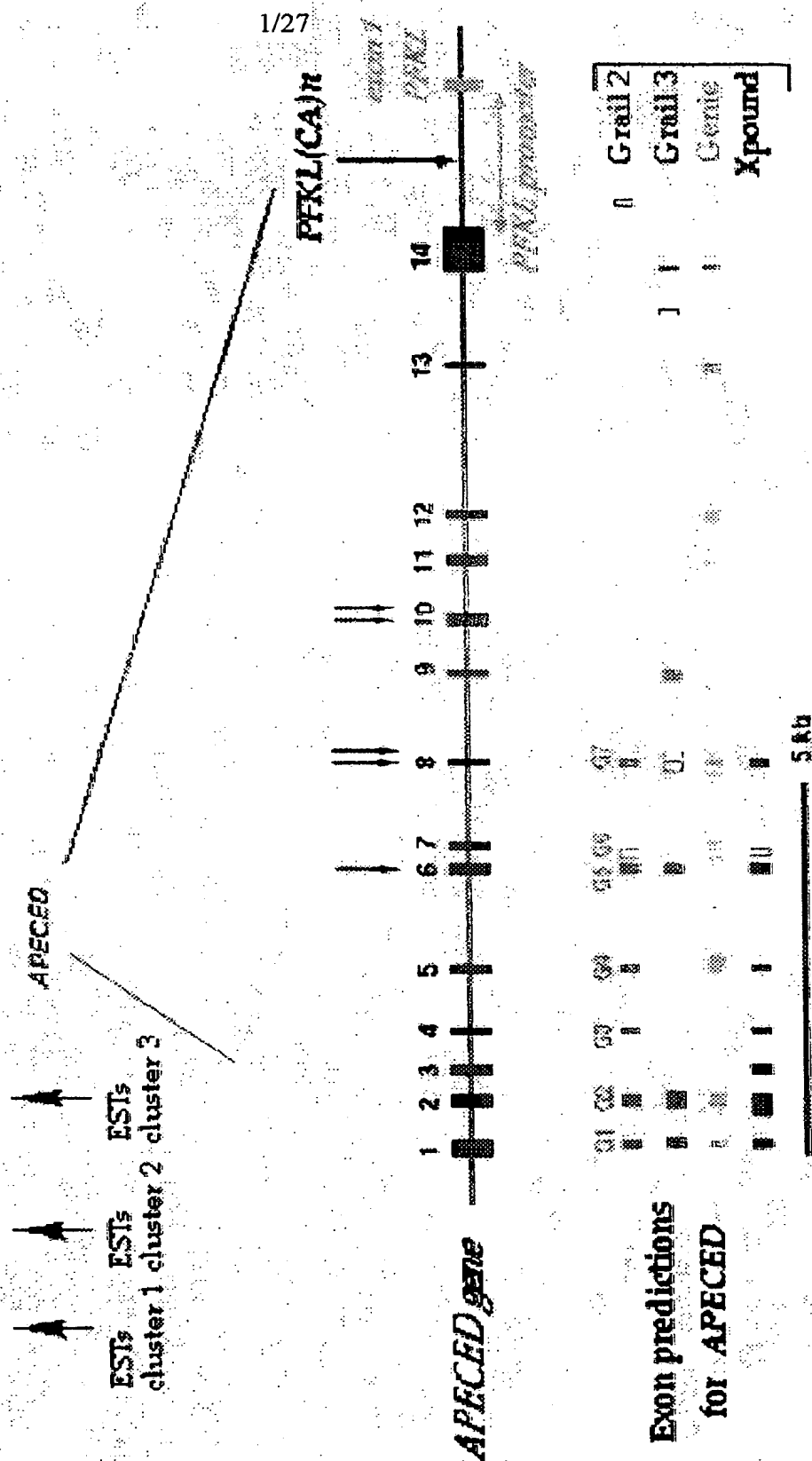


Figure 1B

005020" 56560560

cDNA B1-1

Q21D1

Q11D11



Figure 1C

3/27

1 cgggagcagccggcgaggagggcccccacagcccccgccgggagccccgagggccaagcgagg 60
 61 gctgccagtgtcccgagcccccacccgctccgccccccagcccccggtccccggcccccccc 120
 121 atggcgacggacggcggtacgcccggcttctgagggctgcaccgcacgagagatcgcggtg 180
 1 M A T D A A L R R L L R L H R T E I A V 20
 181 gccgtgacagcgccttcccactgctgcacgctggtgctgaccacgacgtggtccccgag 240
 21 A V D S A F P L L H A L A D H D V V P E 40
 241 gacaagtctcaggagacgcttcatctgaaggaaaaggaggctgccccccagggccttccac 300
 41 D K F Q E T L H L K E K E G C P Q A F H 60
 301 gccctcctgtcctggctgacctgacctccacagcctccacgacctctggacttctggagggtg 360
 61 A L L S W L L T Q D S T A I L D F W R V 80
 361 ctgttcaaggactacaacctggagcgctatggccggctgcagccccatctctggacagcttc 420
 81 L F K D Y N L E R Y G R L Q P I L D S F 100
 421 cccaaagatgtggacctcagccagccccggaggaggaggaggaagcccccgccgtccccaaag 480
 101 P K D V D L S Q P R K G R K P P A V P K 120
 481 gctttgtaccgcccacccagactccccaccagagggaagggcctcagaagaggctcagct 540
 121 A L V P P P R L P T K R K A S E E A R A 140
 541 gccgagcagcagcctgactccaaggggcaccgcccagccaggtctctcaactgaaggcc 600
 141 A A P A A L T P R G T A S P G S Q L K A 160
 601 aagcccccaagaagccggagagcagcgagcagcagcgcccttccactcgggaacggg 660
 161 K P P K K P E S S A E Q Q R L P L G N G 180
 661 attcagaccatgtcagcttccagtcagagagctgtggccatgtctctccggggacgtcccg 720
 181 I Q T M S A S V Q R A V A M S S G D V P 200

FIGURE 2A

721	ggagcccgagggccgtggaggatcctcatccagcaggtgtttgagtcaggcggtcc	780
201	G A R G A V E G I L I Q Q V F E S G G S	220
781	aagaagtgcacccaggttggtgggagttctacactcccagcaagttcgaagactccggc	840
221	K K C I Q V G G E F Y T P S K F E D S G	240
841	agtgggaagaacaaggcccgacagcagcagtgggcccgaaagcctctggttcgagccaaggga	900
241	S G K N K A R S S S S G P K P L V R A K G	260
901	gcccagggcgctgcccccggtggaggtgaggtaggctaggctggccagcaggggcagcgttccc	960
261	A Q G A A P G G G E A R L G Q Q G S V P	280
961	gcccccttgccccctccagtgacccccagctccaccagaagaatgaggacgagtggtgcc	1020
281	A P L A L P S D P Q L H Q K N E D E C A	300
1021	gtgtgtcgggacggcgaggatcctgtgtgacggctgcctcgggccttccacctg	1080
301	V C R D G G E L I C C D G C P R A F H L	320
1081	gcctgcctgtccccctccgctccgggagatccccagtgaggacctggaggtgctccagctgc	1140
321	A C L S P P L R E I P S G T W R C S S C	340
1141	ctgcaggcaacagtcacaggaggtgcagccccgggcagagggagcccccgagagcca	1200
341	L Q A T V Q E V Q P R A E E P R P Q E P	360
1201	cccgtaggagaccccgctcccccggttaggtcgcgaggagaggttaagaggtcca	1260
361	P V E T P L P P G L R S A G E E V R G P	380
1261	cctgggaacccctagccggcatggacacgactcttgtctacaagcacctgcccggctccg	1320
381	P G E P L A G M D T T L V Y K H L P A P	400
1321	ccttctgcagcccccgctgccagggtggactcctcctcgccctgcacccccctactgtgtg	1380
401	P S A A P L P G L D S S A L H P L L C V	420

FIGURE 2A (cont.)

1381	ggtcctgagggcagcagaacctggctcctgggtgcgcgttgcgggtgtgcggagatggg	1440
421	G P E G Q Q N L A P G A R C G V C G D G	440
1441	acggacgtgctgcggtgtactcaactgcgcgctgccttccactggcgtgccacttccca	1500
441	T D V L R C T H C A A A F H W R C H F P	460
1501	gccggcacctcccgcccgggacgggacctgcgtgcagatcctgctcaggagacgtgacc	1560
461	A G T S R P G T G L R C R S C S G D V T	480
1561	ccagccccctgtggaggggtgtggccccccagccccgcctggccccctggccctgcc	1620
481	P A P V E G V L A P S P A R L A P G P A	500
1621	aaggatgacactgccagtcacgagccgctctgcacagggatgacctggagtccttctg	1680
501	K D D T A S H E P A L H R D D L E S L L	520
1681	agcagcacacaccttcgatggcatcctgcagtgggccatccagagcatggccccctccg	1740
521	S E H T F D G I L Q W A I Q S M A R P A	540
1741	gcccccttccccctctgacccccagatggccgggacatgcagctctgatgagagtgctg	1800
541	A P F P S	546
1801	agaaggacacctcctcctcagtcctggaagccggccggtggatcaagaagggacag	1860
1861	cgccacctctgtcagtgctcggtgtaaacagctctgtgtttcttggggacaccagccat	1920
1921	catgtgcctggaaattaaacctgccccacttcttactctggaagtcctccggagccctc	1980
1981	tccttgctggtgacctaataataaaaaattagctgggtgtgtgggtgggtgcctg	2040
2041	taatcccagctacatgggagcctgaggcatgagaatcacttgaactcgggaggtggaggt	2100
2101	tgcagtgagctgagattggccactgcactccagctgtgtgtggcgaagagtgcactccgt	2160
2161	ctcaaaaaacaaaaaa ^{pl-1} ccacataacataaaatttatcatctcgaccacttttcagt	2220
2221	tcagtggcattcacatctcatgtaa	2245

FIGURE 2A (cont.')

6/27

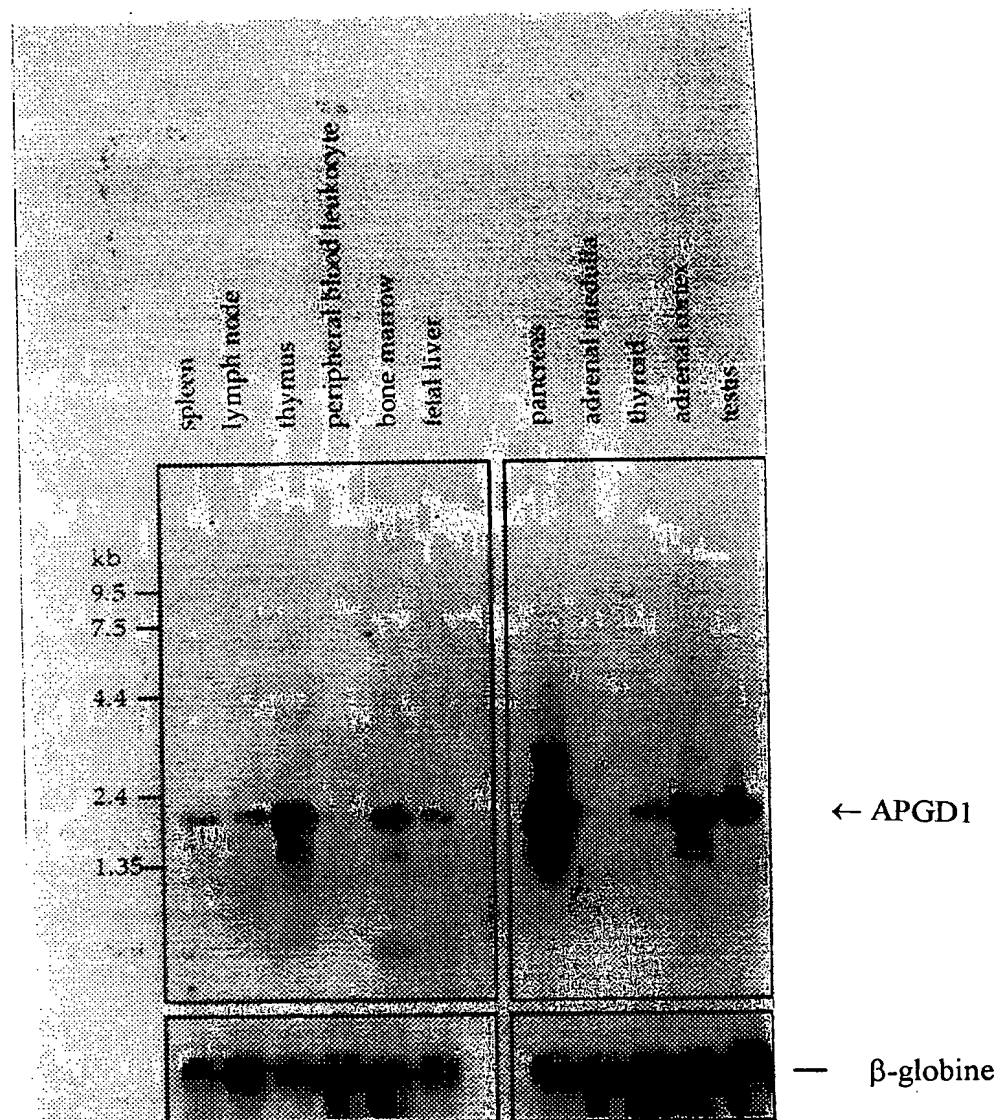
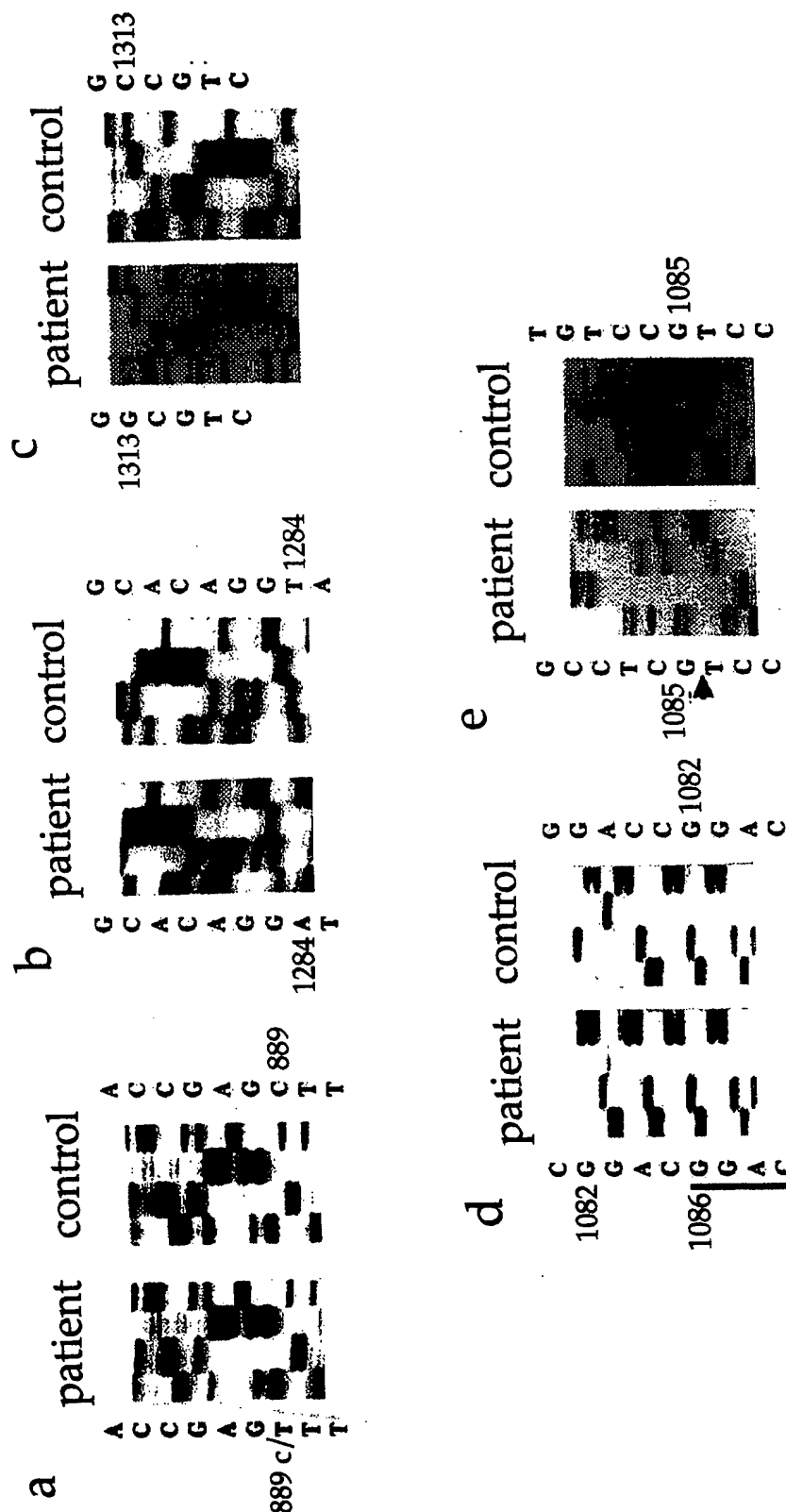


Figure 2B



The sequence lanes appear from left to right, as C, A, T, and G

Figure 3

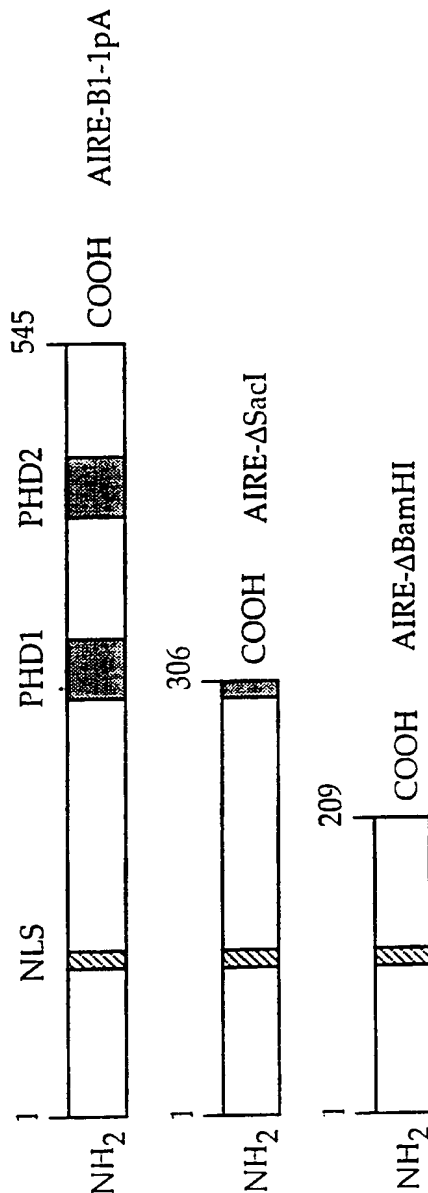


Figure 4

9/27

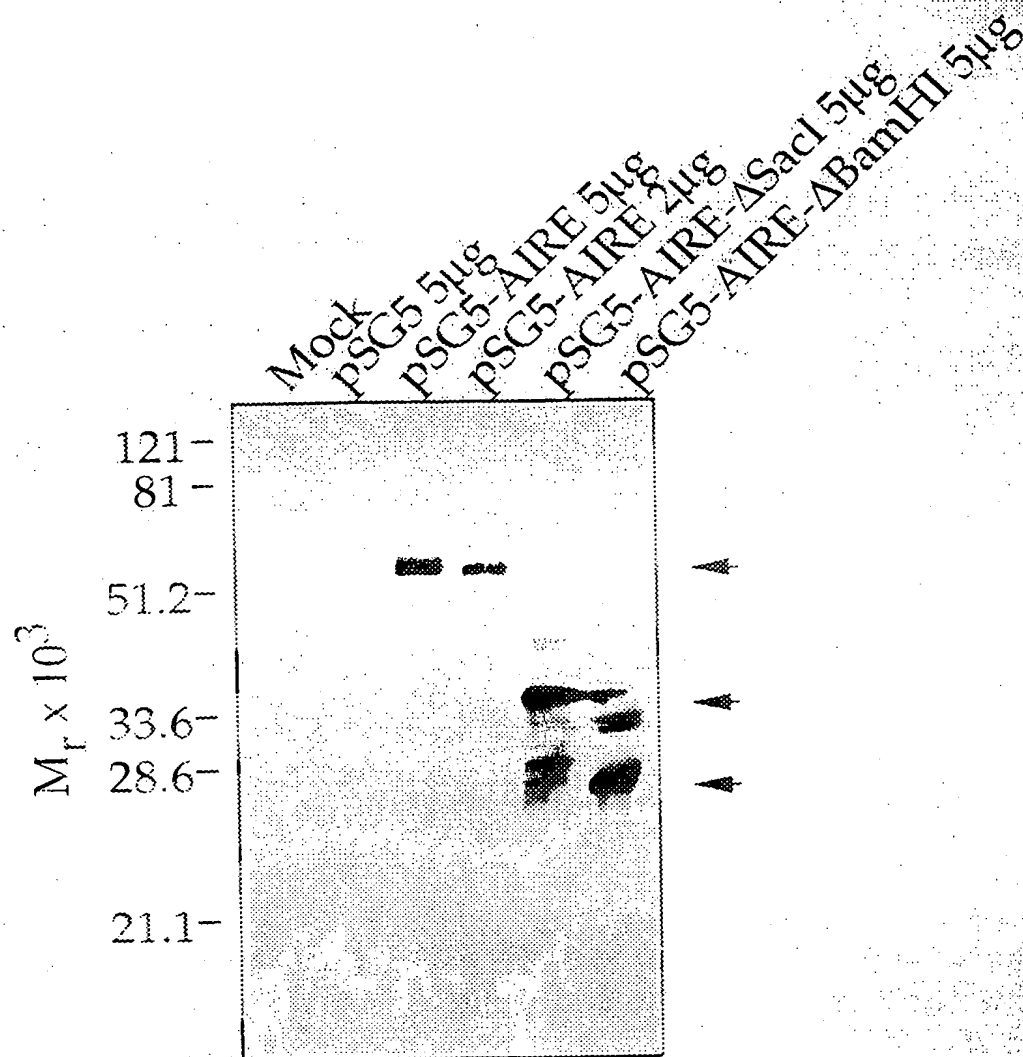


Figure 5

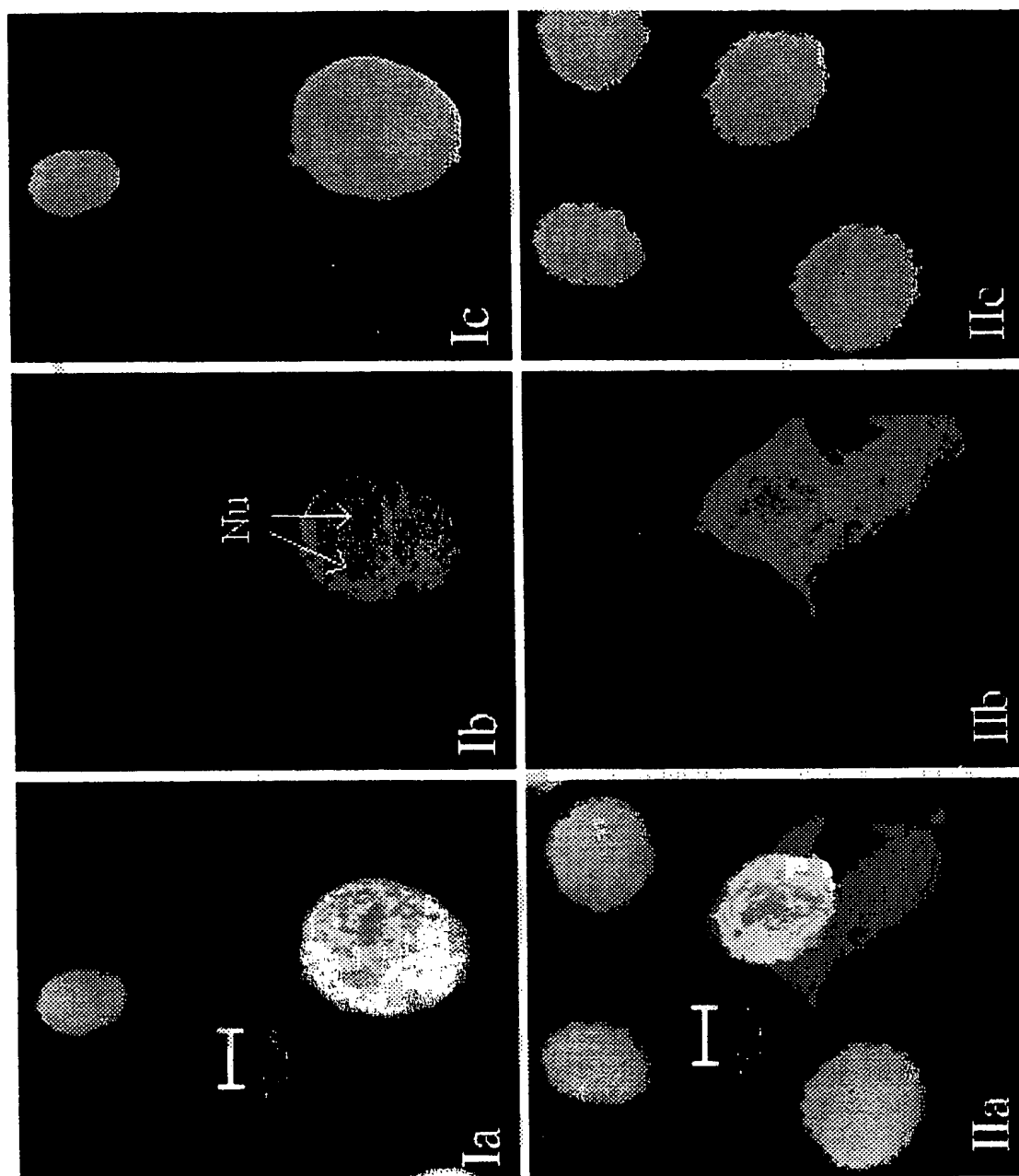
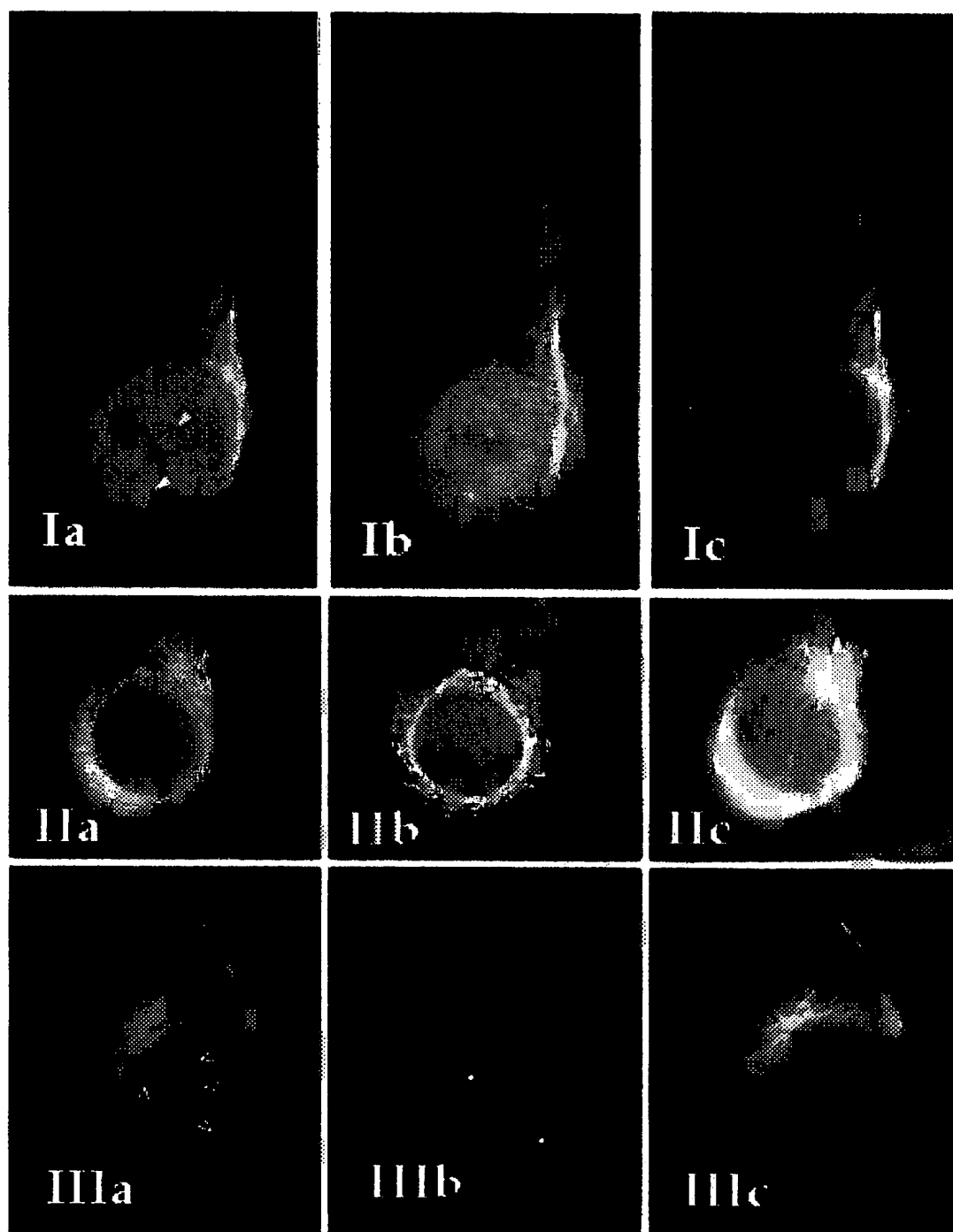


Figure 6

11/27

**Figure 7**

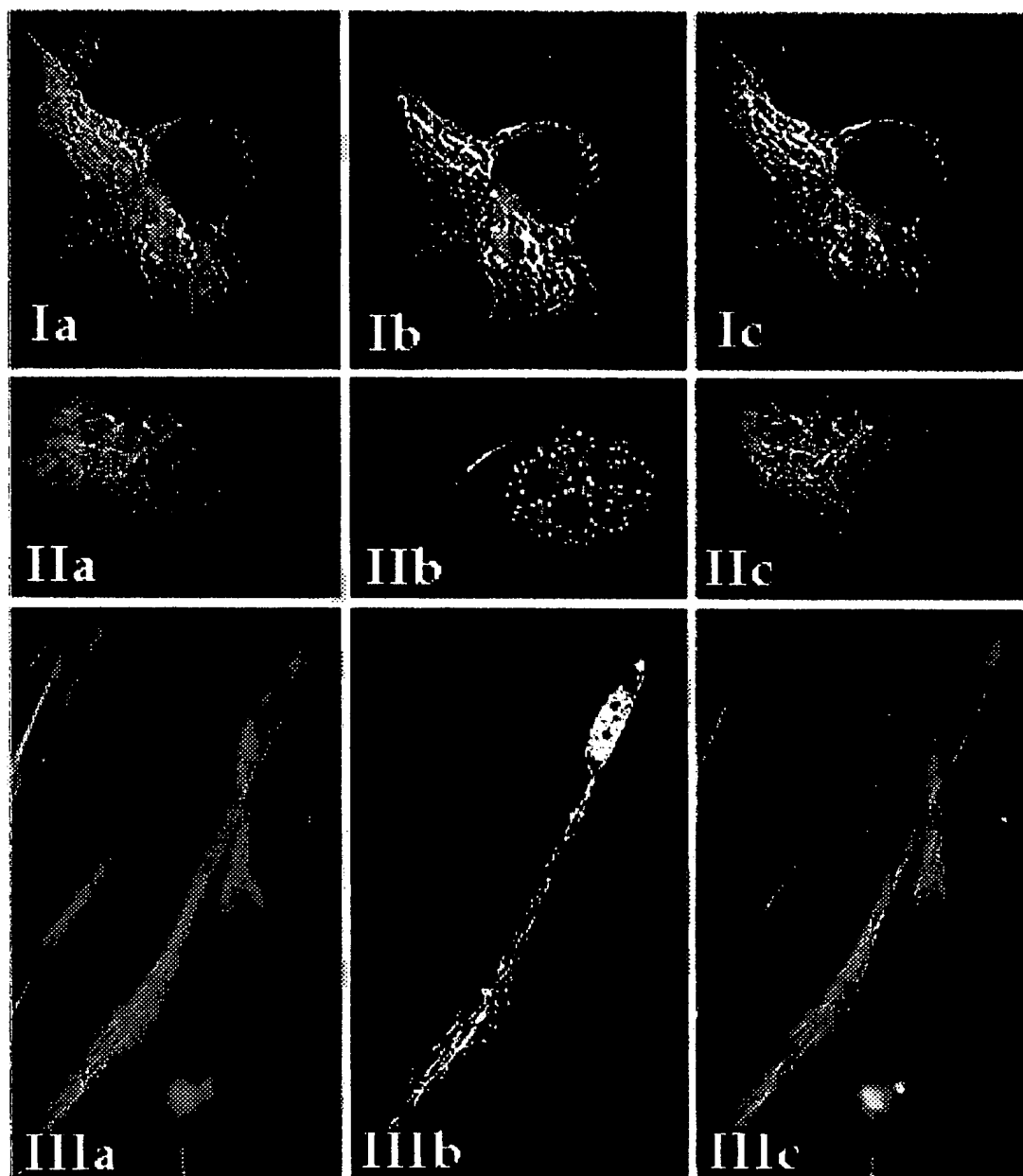


Figure 8

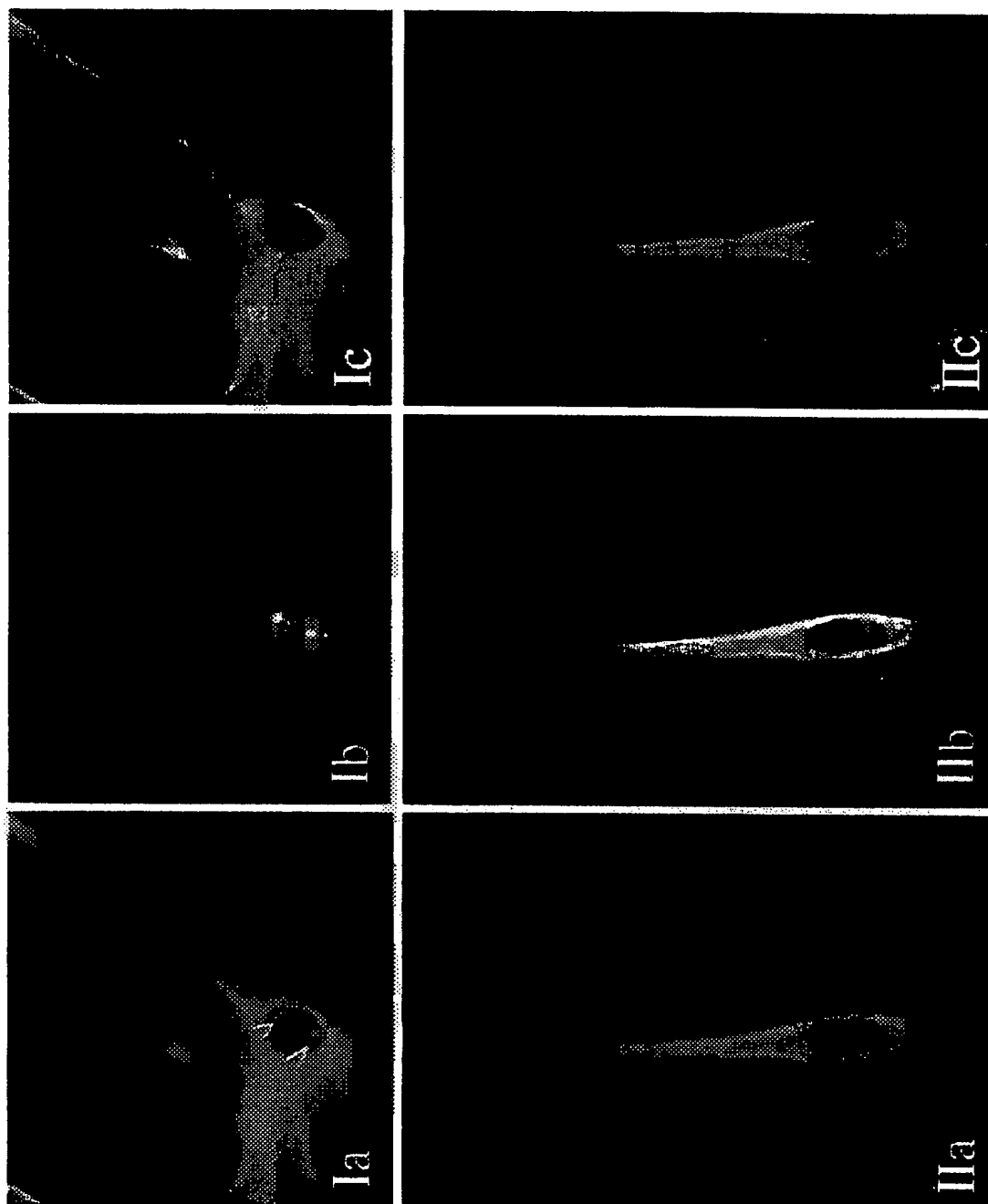


Figure 9

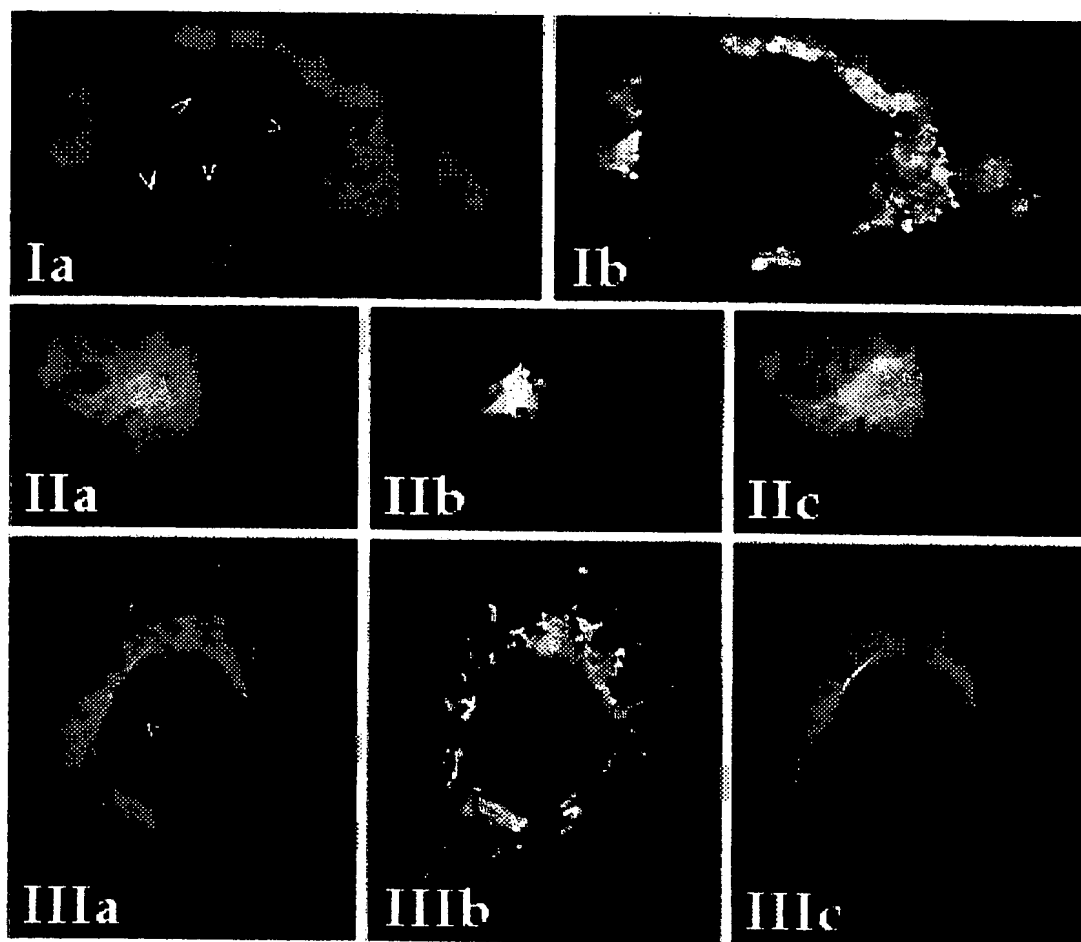


Figure 10

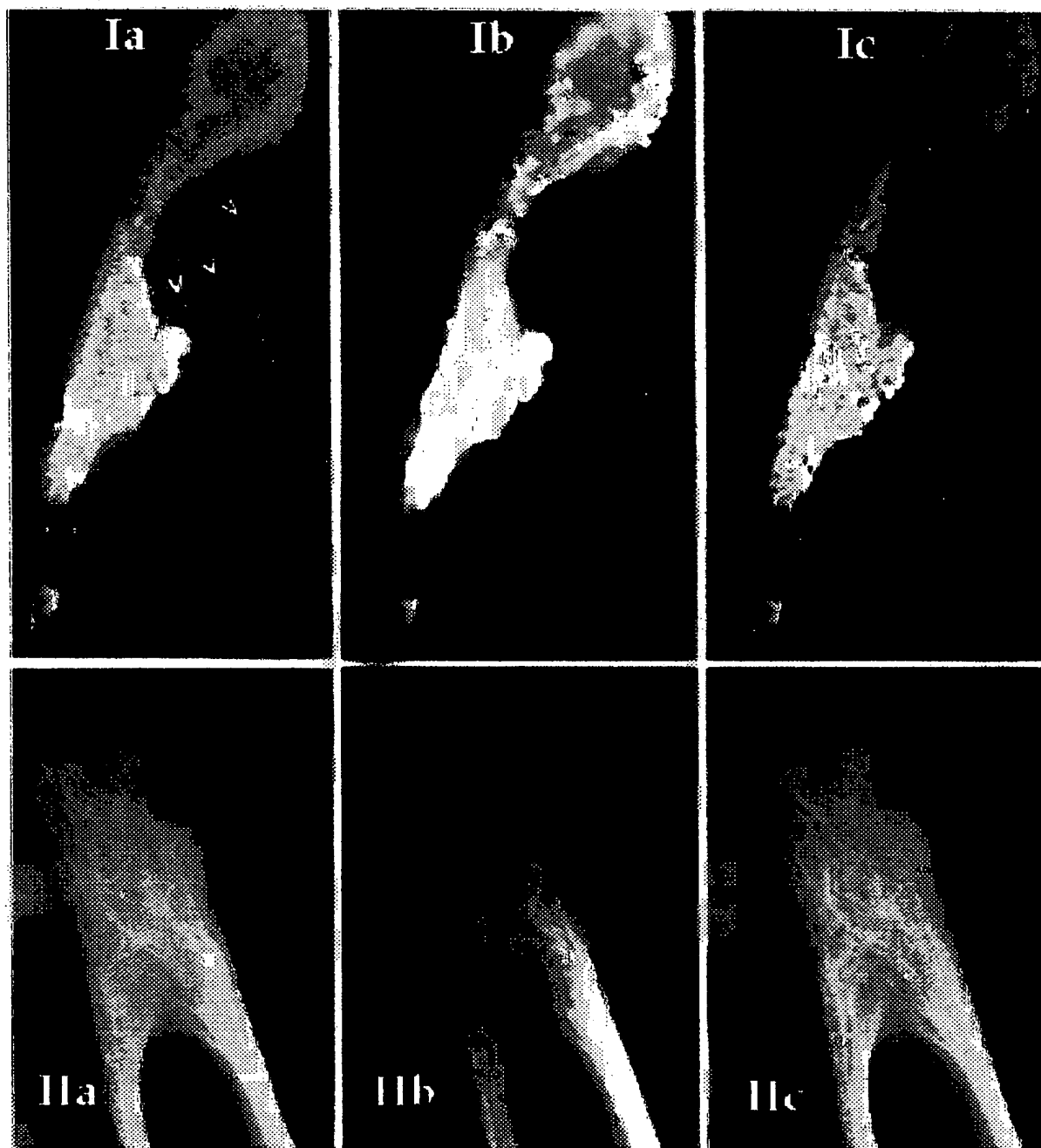
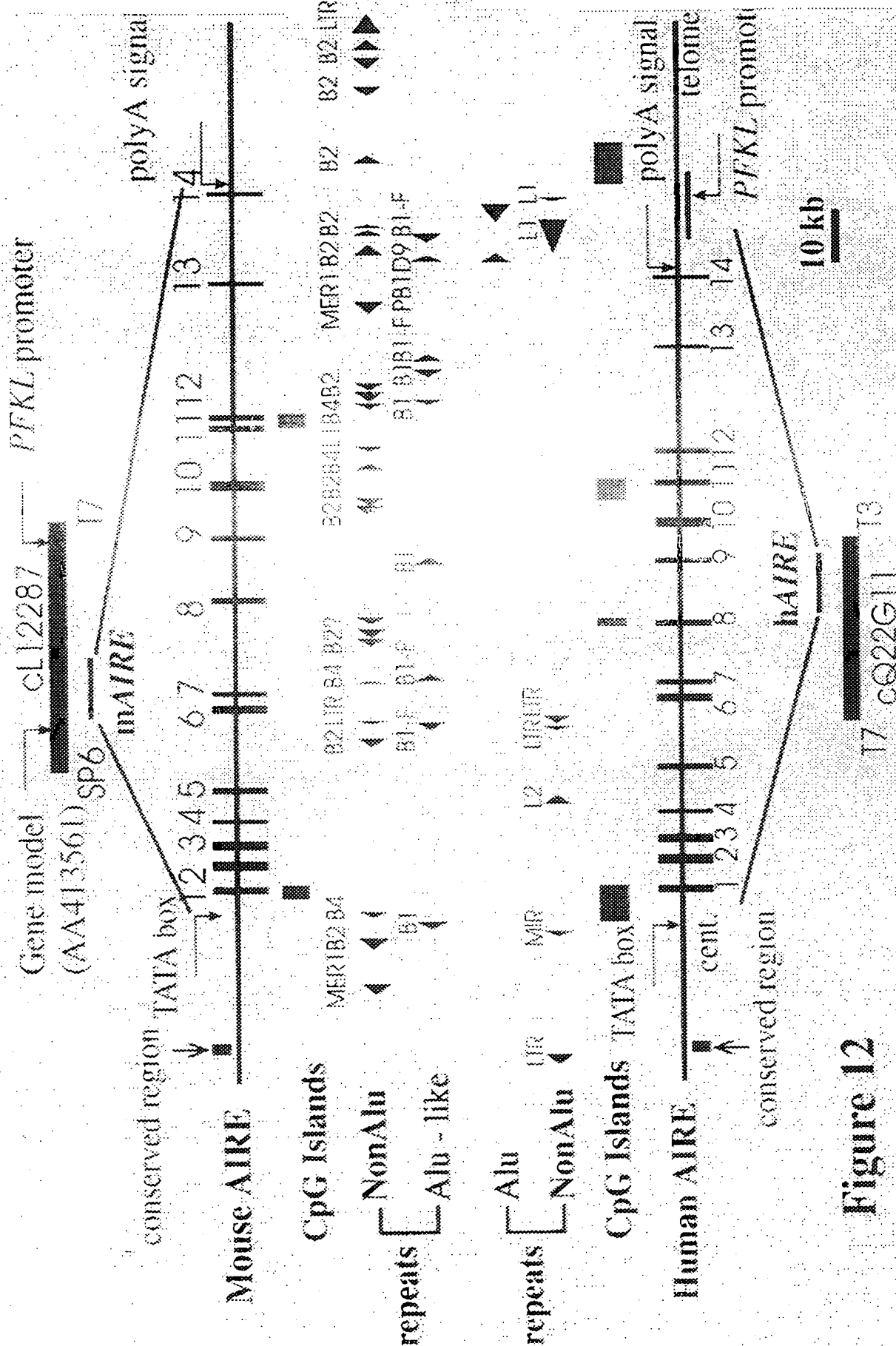


Figure 11

16/27



Index

17/27

HSAJ9610 (horizontal) vs. AF073797 (vertical)

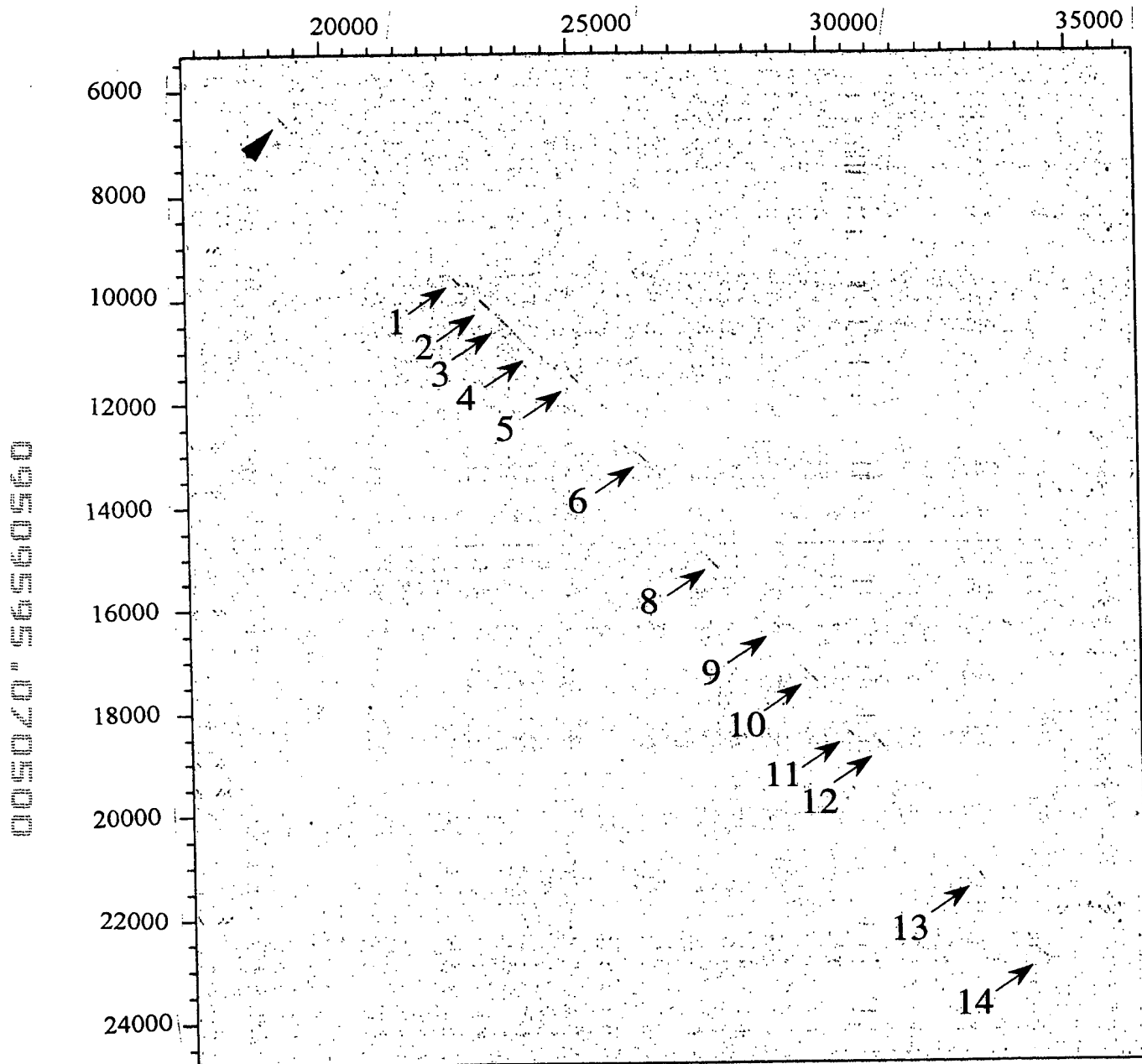


Figure 13A

005020" 56560560

6486 6575
LAIRE GTGTGGACTG TCACGGAAAC CCCACGCTGT GATGGAAAGT CCAAAATTCT ACAGGAGTCT TTCTGTTGAT CTCCAGTCAG AGGCTGGGG
LAIRE AAGGGGCTGG TGTGGAAAGC CCCACGGCAT GGTGGAAAGT CCGAAATTCT ACAGGGGCTT CTTGTTAAA CCTCCATGCA AGAGGCTGG
19186 19275
ons. --G-GG---G T---G-AA-C CCC---G--T G-TGGAAAGT CC-AAATTCT ACAGG-G-CT -T-TGTT-A- C--C--T--- AG-----GGG

Figure 13B

19/27

	10	30	50	
1	ATGGCAGGTGGGGATGGAATGCTACGCCGTCTGCTGAGGCTGCACCGCACCGAGATCGCG			60
1	M A G G D G M L R R L L R L H R T E I A			20
	70	90	110	
61	GTGGCCATAGACAGTGCCTTTCCGCTGCTGCATGCTCTAGCCGACCACGACGTGGTCCCT			120
21	V A I D S A F P L L H A L A D H D V V P			40
	130	150	170	
121	GAGGACAAGTTCCAGGAGACGCTCCGTCTGAAGGAGAAGGAAGGCTGCCCCAGGCCTTC			180
41	E D K F Q E T L R L K E K E G C P Q A F			60
	190	210	230	
181	CACGCCCTGCTGTCCTGGCTCCTGACCCGGGACAGTGGGGCCATCCTGGATTTCTGGAGG			240
61	H A L L S W L L T R D S G A I L D F W R			80
	250	270	290	
241	ATTCTCTTTAAGGACTACAATCTGGAGCGGTACAGCCGCCTGCATAGCATCCTGGACGGC			300
81	I L F K D Y N L E R Y S R L H S I L D G			100
	310	330	350	
301	TTCCCAAAGATGTGGACCTAAACCAGTCCCGGAAAGGGAGAAAGCCCCCTTGCTGGTCCC			360
101	F P K D V D L N Q S R K G R K P L A G P			120
	370	390	410	
361	AAGGCCGCGGTACTGCCACCCAGACCCCCACCAAGAGAAAAGCACTGGAGGAGCCTCGA			420
121	K A A V L P P R P P T K R K A L E E P R			140
	430	450	470	
421	GCCACCCCAACAGCAACTCTGGCCTCAAAGAGCGTCTCCAGCCCAGGCTCCACCTGAAG			480
141	A T P P A T L A S K S V S S P G S H L K			160
	490	510	530	
481	ACTAAGCCCCCTAAGAAGCCAGATGGCAACTTGGAGTCAACAGCACCTTCCTCTTGGAAC			540
161	T K P P K K P D G N L E S Q H L P L G N			180
	550	570	590	
541	GGAATTCAGACCATGGCAGCTTCTGTCCAGAGAGCTGTGACCGTGGCCTCTGGGGATGTT			600
181	G I Q T M A A S V Q R A V T V A S G D V			200
	610	630	650	
601	CCAGGAACCCGAGGGGCCGTGGAAGGGATCCTTATCCAGCAGGTGTTTGAGTCAGGAAGA			660
201	P G T R G A V E G I L I Q Q V F E S G R			220
	670	690	710	
661	TCCAAGAAGTGCATTACAGGTTGGGGGAGAGTTTTATACACCCAACAAGTTCGAAGACCCC			720
221	S K K C I Q V G G E F Y T P N K F E D P			240

Figure 14A

20/27

	730	750	770	
721	AGTGGCAATTTGAAGAACAAGGCCCGGAGTGGTAGCAGCCTAAAGCCAGTGGTCCGAGCC			780
241	S G N L K N K A R S G S S L K P V V R A			260
	790	810	830	
781	AAGGGAGCCCAGGTCACTATACCTGGTAGAGATGAGCAGAAAGTGGGCCAGCAGTGTGGG			840
261	K G A Q V T I P G R D E Q K V G Q Q C G			280
	850	870	890	
841	GTTCCCTCCCTTCCATCCCTCCCCAGTGAGCCCCAGGTTAACCAGAAGAACGAGGATGAG			900
281	V P P L P S L P S E P Q V N Q K N E D E			300
	910	930	950	
901	TGTGCCGTGTGCCACGACGGAGGTGAGCTCATCTGTTGTGACGGCTGTCCCCGGGCCTTC			960
301	C A V C H D G G E L I C C D G C P R A F			320
	970	990	1010	
961	CACCTGGCTTGCCTGTCCCCACCTCTGCAGGAGATCCCCAGTGGCCTCTGGAGATGCTCC			1020
321	H L A C L S P P L Q E I P S G L W R C S			340
	1030	1050	1070	
1021	TGCTGCCTCCAGGGCAGAGTCCAACAGAACCTGTCCCAGCCTGAGGTGTCCAGGCCCCCG			1080
341	C C L Q G R V Q Q N L S Q P E V S R P P			360
	1090	1110	1130	
1081	GAGCTACCTGCAGAGACCCCGATCCTCGTGGGACTGAGGTCAGCTTCAGAGAAAACCAGG			1140
361	E L P A E T P I L V G L R S A S E K T R			380
	1150	1170	1190	
1141	GGCCCATCCAGGGAGCTCAAAGCCAGCTCTGATGCTGCTGTACATATGTGAACCTGCTG			1200
381	G P S R E L K A S S D A A V T Y V N L L			400
	1210	1230	1250	
1201	GCCCCGCACCTGCAGCTCCTCTGCTGGAGCCTTCAGCACTGTGCCCTCTACTGAGTGCT			1260
401	A P H P A A P L L E P S A L C P L L S A			420
	1270	1290	1310	
1261	GGGAATGAGGGGCGGCCAGGTCCAGCACCAAGCGCGCATGCAGTGTGTGTGGCGATGGC			1320
421	G N E G R P G P A P S A R C S V C G D G			440
	1330	1350	1370	
1321	ACCGAGGTGTTGCGGTGTGCACACTGTGCCGCTGCCTTCCACTGGCGCTGCCACTTCCCG			1380
441	T E V L R C A H C A A A F H W R C H F P			460
	1390	1410	1430	
1381	ACGGCCGCCGCCGGCGGGGACCAATCTCCGCTGCAAATCCTGCTCTGCAGACTCGACT			1440
461	T A A A R P G T N L R C K S C S A D S T			480
	1450	1470	1490	

Figure 14B

21/27

1441	CCCACGCCAGGCACACCGGGCGAAGCTGTACCCACCTCTGGGCCCCGTCCAGCACCTGGG	1500
481	P T P G T P G E A V P T S G P R P A P G	500
	1510 1530 1550	
1501	CTTGCCAAGgtagGGGACGACTCTGCTAGTCACGACCCTGTTCTACATAGGGACGACCTG	1560
501	L A K V G D D S A S H D P V L H R D D L	520
	1570 1590 1610	
1561	GAGTCCCTCCTCAATGAGCACTCATTGACGGCATCCTGCAGTGGGCCATCCAGAGCATG	1620
521	E S L L N E H S F D G I L Q W A I Q S M	540
	1630 1650	
1621	TCACGCCCGCTGGCCGAGACACCACCCTTCTCTTCC	1656
541	S R P L A E T P P F S S	552

Figure 14C

22/27

M 1 2 3 4 5 6

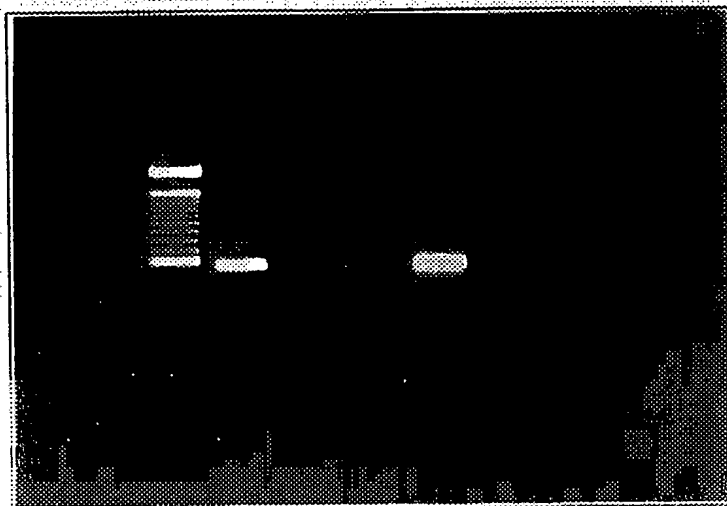


Figure 15

23/27

Human AIRE	-MATDAALRR	LI	RLHRTEIA	VAVDSAFPLL	HALADHDVVP	EDKFQETLHL
Mouse AIRE	MAGGDGMLRR	LI	RLHRTEIA	VAIDSAFPLL	HALADHDVVP	EDKFQETLRL
Consensus	----D--LRR	LI	RLHRTEIA	VA-DSAFPLL	HALADHDVVP	EDKFQETL-L
	51					100
Human AIRE	KEKEGCPQAF	HALLSWLLTQ	DSTAILDFWR	VLFKDYNLER	YGRLOPILDS	
Mouse AIRE	KEKEGCPQAF	HALLSWLLTR	DSGAILDFWR	ILFKDYNLER	YSRLHSILDG	
Consensus	KEKEGCPQAF	HALLSWLLT-	DS-AILDFWR	-LFKDYNLER	Y-RL--ILD-	
	101					150
Human AIRE	FPKDVDLSQP	RKGRKPPAVP	KALVPPPRLP	TKRKASEEAR	AAAPAAALTPR	
Mouse AIRE	FPKDVDLNQS	RKGRKPLAGP	KA AVLPPRPP	TKRKALEEPR	ATPPATLASK	
Consensus	FPKDVDL-Q-	<u>RKGRKP</u> -A-P	KA-V-PPR-P	<u>TKRKA</u> -EE-R	A--PA-L---	
	151					200
Human AIRE	GTASPGSQLK	AKPPKKPESS	AEQQRLLPLGN	GIQTMSASVQ	RAVAMSSGDV	
Mouse AIRE	SVSSPGSHLK	TKPPKKPDGN	LESQHLPLGN	GIQTMAASVQ	RAVTVASGDV	
Consensus	---SPGS-LK	-KPPKKP---	-E-Q-LPLGN	GIQTM-ASVQ	RAV---SGDV	
	201					250
Human AIRE	PGARGAVEGI	LIQQVFESGG	SKKCIQVGGE	FYTPSKFED.	SGSGKNKARS	
Mouse AIRE	PGTRGAVEGI	LIQQVFESGR	SKKCIQVGGE	FYTPNKFEDP	SGNLKNKARS	
Consensus	PG-RGAVEGI	LIQQVFESG-	SKKCIQVGGE	FYTP-KFED-	SG--KNKARS	
	251					300
Human AIRE	SSGPKPLVRA	KGAQGAAPGG	GEARLGQQGS	VPAPLALPSD	PQLHQKNEDE	
Mouse AIRE	GSSLKPVVRA	KGAQVTIPGR	DEQKVGQQCG	VPPLPSLPSE	PQVNQKNEDE	
Consensus	-S--KP-VRA	<u>KGAQ</u> ---PG-	-E---GQQ--	VP----LPS-	PQ--QKNEDE	
	301					350
Human AIRE	CAVCRDGGEL	ICCDGCPRAF	HLACLSPPLR	EIPSGTWRCs	SCLQATVQEV	
Mouse AIRE	CAVCHDGGEL	ICCDGCPRAF	HLACLSPPLO	EIPSGLWRCS	CCLQGRVQQN	
Consensus	CAVC-DGGEL	ICCDGCPRAF	HLACLSPPL-	EIPSG-WRCS	-CLQ--VQ--	
	351					400
Human AIRE	QPRAEEPRPQ	EPPVETPLPP	GLRSAGEEVR	GPPGEPLAGM	DTTLVYKHLp	
Mouse AIRE	LSQPEVSRPP	ELPAETPILV	GLRSASEKTR	GPSRELKASS	DAAVTYVNNL	
Consensus	----E--RP-	E-P-ETP---	GLRSA-E--R	GP--E--A--	D----Y--L-	
	401					450
Human AIRE	APPSAAPLPG	LDSSALHPLL	CVGPEGQQNL	APGARCGVCG	DGTDVLRCTH	
Mouse AIRE	APHPAAPL..	LEPSALCPLL	SAGNEGRPGP	APSARCSVCG	DGTEVLRCAH	
Consensus	AP--AAPL--	L--SAL-PLL	--G-EG----	AP-ARC-VCG	DGT-VLRC-H	
	451					500
Human AIRE	CAAAFHWRCR	FPAGTSRPGT	GLRCRSCSGD	VTPAP.VEGV	LAP.SPAPLA	
Mouse AIRE	CAAAFHWRCR	FPTAAARPGT	NLRCKSCSAD	STPTPGTPGE	AVPTSGPRPA	
Consensus	CAAAFHWRCR	FP----RPGT	-LRC-SCS-D	-TP-P---G-	--P-S--R-A	
	501					550
Human AIRE	PGPAK..DDT	ASHEPALHRD	DLESLLSEHT	FDGILQWAIQ	SMARPAAPFP	
Mouse AIRE	PGLAKVGDDs	ASHDPVLHRD	DLESLLNEHS	FDGILQWAIQ	SMSRPLAETP	
Consensus	PG-AK--DD-	ASH-P-LHRD	DLESLL-EH-	FDGILQWAIQ	SM-RP-A--P	
	551					
Human AIRE	S---					
Mouse AIRE	PFSS					
Consensus	----					

Figure 16

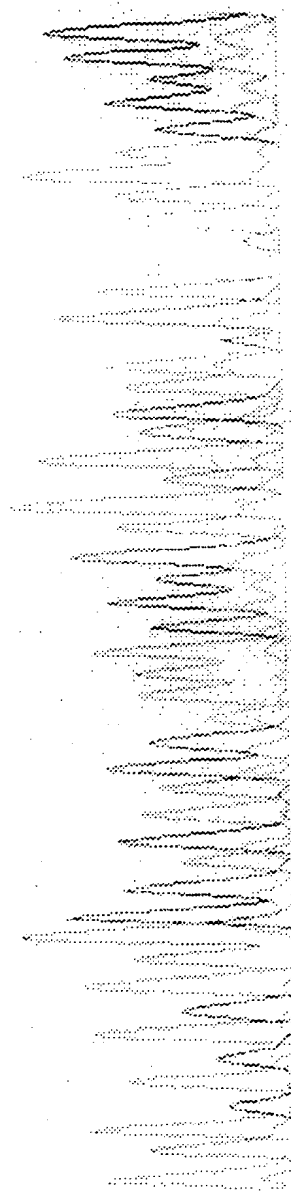
Exon II

Δ Exon 10

Exon 9

R A S P A P Q P T E A P L R P

3'-TCCGCGCTTCGGTGCTGGAGCAGGGGTCATCTGGGCGGT-GTTGGGG-5'



W
D
C
G
L

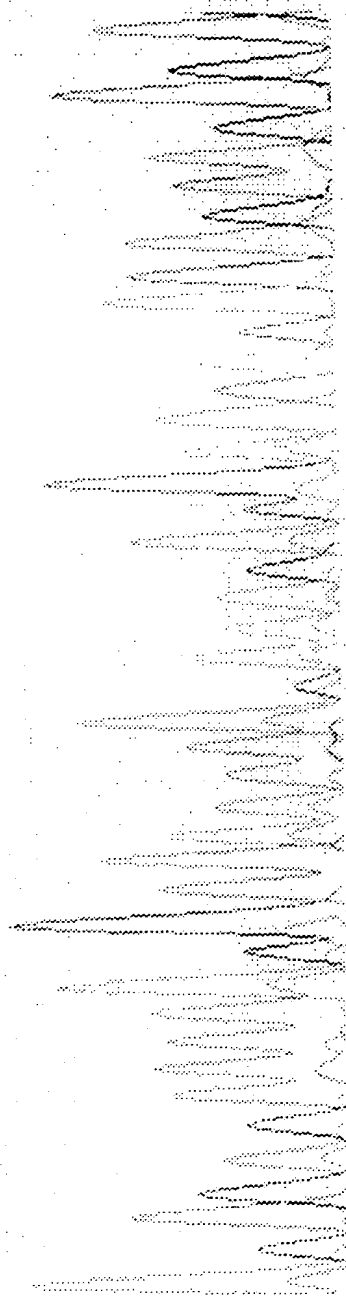
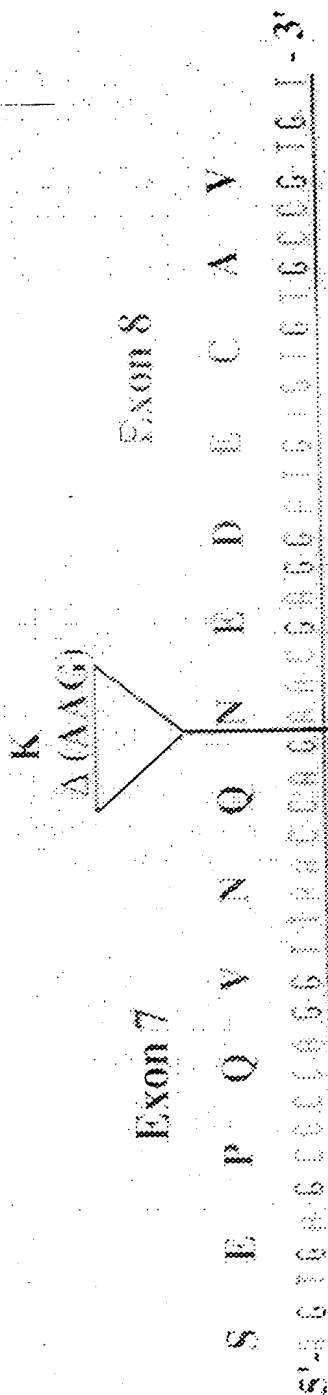
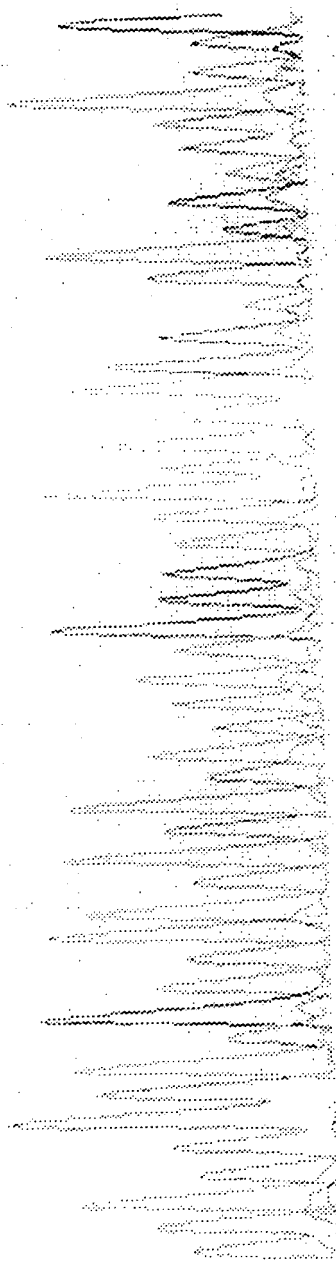


FIGURE 17B

INTRODUCTION

Y K Q B D R G Q A C K A R V V



0176

27/27

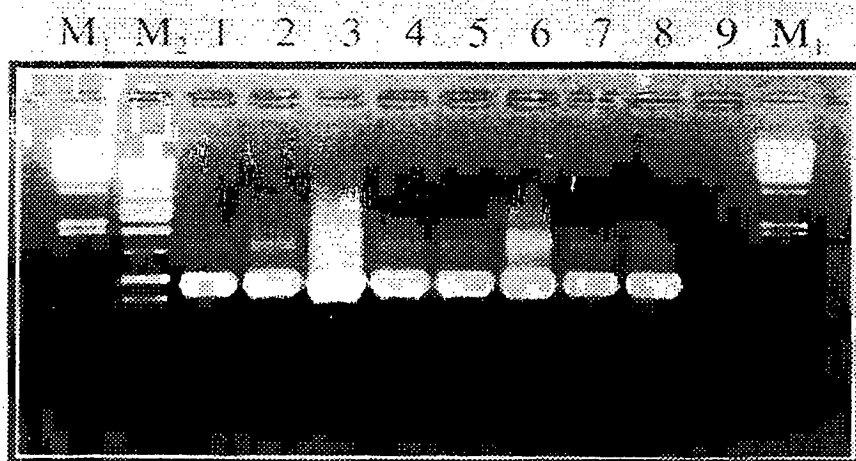


Figure 18